DOCUMENT RESUME

ED 076 201 LI 004 306

AUTHOR Spyers-Duran, Peter
TITLE Secondary Access Storage of Books in Small and Medium-Sized Academic Libraries: A Proposal for an Experimental Model.
INSTITUTION Florida Atlantic Univ., Boca Raton.
PUB DATE May 73
NOTE 22p.; (10 References)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Books; *College Libraries; Library Collections; *Library Facilities; Library Materials; Library Planning; Research Design; *Storage; *University Libraries; Use Studies
IDENTIFIERS *Book Storage

ABSTRACT

The steady growth of the academic libraries reflects a true increase in the body of human knowledge. The average academic library either now has, or will have within the next few years, a severe space problem related to housing of books. The seemingly limitless growth of space needs creates problems compounded by rising costs of buildings and equipment. Effective future planning of academic libraries would be greatly facilitated if there were some theoretical bases which could be used to predict the probable distribution of the future use of library materials in subject matters that have curricular significance on a given campus. This paper proposes to design a method by which small and medium-sized academic libraries can determine the relationship between usage and age of monographic material, and the application of this relationship to consideration of the possible establishment of a less expensive secondary access storage facility. Since studies of this magnitude are expensive, a pilot study at the Florida Atlantic University campus is proposed to test for design weaknesses and validity. The results should be useful to other academic libraries in terms of research design and other appropriate applications. (Author/NH)

FILMED FROM BEST AVAILABLE COPY

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED 076201

SECONDARY ACCESS STORAGE OF BOOKS IN SMALL AND MEDIUM-SIZED ACADEMIC LIBRARIES: A PROPOSAL FOR AN EXPERIMENTAL MODEL

by

Peter Spyers-Duran

May 1973

LI 004 306

I. INTRODUCTION

The steady growth of the academic libraries reflects a true increase in the body of human knowledge. The average academic library either now has, or will have within the next few years, a severe space problem related to housing of books. The seemingly limitless growth of space needs creates problems compounded by rising costs of buildings and equipment.

Effective future planning of academic libraries would be greatly facilitated if there were some theoretical bases which could be used to predict the probable distribution of the future use of library materials in subject matters that have curricular significance on a given campus.

This paper proposes to design a method by which small and medium-sized academic libraries can determine the relationship between usage and age of monographic material, and the application of this relationship to consideration of the possible establishment of a less expensive secondary access storage facility.

Since studies of this magnitude are expensive, a pilot study at the Florida Atlantic University campus is proposed to test for design weaknesses and validity. The results are expected to be useful to other academic libraries in terms of research design and other appropriate applications.

II. FACTORS GENERATING SPACE NEEDS and STUDIES OF USAGE AND STORAGE

Space needs in libraries are generated by several factors that must be recognized.

The first factor is the fast growing collections. There

were twice as many books published in 1972 than in the mid-1950's. As a consequence, libraries add twice as many works to their collections annually than they did twenty years ago. In general, library planners agree that a typical academic library will dedicate 30 to 40 per cent of its net assignable floor space to stacks.

The second space-need factor is the size and level of service expectation of the student body and faculty. The enrollments have increased on every campus year after year. The student arriving from the junior college, or from the high school, expects to find a degree of sophistication in libraries that was unheard of ten to fifteen years ago. The American Library Association Standards for College and University Libraries recommend seating for 25 per cent of the student body. Libraries usually fail to meet this level of seating, but we can estimate that 35 to 45 per cent of a typical library building is dedicated to study areas.

The third factor generating space need is the service areas. The demand for more, better and new services has created a condition wherein library administrators had to consider allocating greater than the traditional 15 to 20 per cent of net assignable floor space to this function.

While examining the possible solutions and alternatives to housing collections, we must remind ourselves that efficient access is essential to all users. This implies that any economical solution will be acceptable to the entire clientele only if the efficient access includes both suitable physical arrangements and a continuing systematic acquisition of new materials.

In cases of past space shortages, library administrators have evaluated the alternatives available and allocated available space for the highest local priority at the expense of all other

space needs. Typically, when the stack load nears capacity, study or service space will be converted into additional stack space. The opposite is known to be true only in cases where the library has gained new space for its needs.

In the past, libraries have responded to some aspects of collection growth that can be categorized in four groups:

1) Subdivision of the library into smaller units.
2) Extensive use of microformat.
3) Cooperative collection development among libraries.
4) Secondary level and compact storage of infrequently used materials.

How successful have these methods been in relieving space needs?

First, the subdivision of the library into smaller units has been practiced for years. It has provided temporary relief; it has also resulted in scattering the collection, causing recognized disadvantages. We have seen rapid growth of departmental libraries which filled up fast with no room to expand. Many large universities have attempted to pull back their decentralized library units in recent years, such as University of Chicago, Michigan State and University of Indiana.

The second alternative - the extensive use of microformat - has been applied in many institutions. This measure has already saved thousands of square feet of storage space in many libraries. There are only a few libraries which have not replaced their old bound newspapers and some back issues of journals with more convenient microfilm. In spite of the advances made in this direction, the microformats have proven to be unsuccessful as the total remedy to space utilization in libraries. There are at least three

major reasons to this effect: (a) there are limits to what is available in microformat, (b) micropublishing lags in time behind hard copy publishing industry and (c) there is a definite user preference for hard copy material.

The third alternative lies in the possible development of strong regional supporting collections to which universities may turn with formal assurances of availability when needed, and in certain categories of infrequently used material. The limitation of this alternative is that it saves space only in an area of peripheral interest to a given campus. It is unlikely that a library will have extensive holdings in subjects of peripheral interests, hence the impact on space will probably be negligible.

The fourth alternative suggests that not all library material needs to occupy primary building space. Infrequently used material may possibly be stored in a secondary-level-of-storage facility without placing the students and faculty at a disadvantage.

Since none of the alternatives can be considered as the only salvation to space needs on campuses, it is assumed that findings will be evaluated in light of local needs and circumstances.

This study focuses on the fourth alternative that suggests a secondary storage for some categories of library material when appropriate local conditions exist. For purposes of this study, the primary storage space is identified as an area open to the students and faculty which is also physically an integral part of the library that serves that group. The secondary level of storage is to be identified as a storage facility either on or off campus. Such a facility is neither convenient for use nor suitable for browsing. For these reasons, material wanted from storage is delivered to the primary facility upon demand.

Since secondary storage level works only for material that is infrequently called for, the major question involves testing the methodology that is suitable to identify the least used groups of material in a given library. A closely related interest of the study is the testing of the accuracy of prediction of numbers of withdrawals from the group that has been sent to storage.

III. SURVEY OF LITERATURE RELATING TO SELECTION OF LIBRARY MATERIAL FOR STORAGE

A considerable amount of attention has been given to the problem of storage of library materials. As the space problem reaches a critical stage, the determination to seek solutions to the storage question has been strengthened.

In dealing with the problem of growth, an examination of literature shows two types of solutions that have emerged over the years: (1) development of on-campus storage facilities and (2) development of multi-institutional cooperative storage facilities. The selection of the type of storage facility is part of a larger issue that Cox calls the "total storage decision problem". He raises two basic questions: (a) how will the books be selected for storage and (b) how and where will they be stored.[1]

The basic question of interest to this paper may be phrased as "what are the alternatives for predicting the future use of currently used material?". Can statistical models that consider objective characteristics such as date of publication, language, past use, be helpful? Should the decision be based on the judgement of subject experts? Should it be a combination of both?

[1] Cox, Julius Grady - "Optimum Storage of Library Materials", Lafayette, Purdue University Libraries, 1964, p. 7

The common assumption of the practitioner is that the least used material should be stored. The problem created by this rather simple assumption is vast, since a method of arranging the collection in descending order of probable future use is needed. The professional literature has references to subjective evaluation methodologies.

Cooper writes about her experiences in reducing the size of the Chemistry Library at Columbia University. The collection was reviewed jointly by the library staff and university faculty almost title by title.[2] Again, the object here was to establish which portion of the collection could be satisfactorily housed in the secondary access facility. The weeding method described by Cooper is judged to be too time consuming and economically impractical.

Lister maintains that intellectual weeding policies which require individual judgement and are based on undefined variables usually turn out to be unmanageable projects, especially when collections are large.[3] Lister advocates that criteria for selection to storage should be based on the current or immediate rate of circulation.

Bradford observed as early as 1934 that most of the articles on a given subject are derived from a small proportion of the total journal titles available. Bradford listed the journals

[2]Cooper, Marianne - "Criteria for Weeding Collections", Library Resources & Technical Sciences, 12 339 51, Summer 1968

[3]Lister, Winston C. - "Least Cost Decision Rules for the Selection of Library Materials for Compact Storage", Lafayette, Library Operations Research Project, School of Industrial Engineering, Purdue University 1967

concerned in order of decreasing productivity and, by plotting the logarithms of the cumulative totals of titles against the cumulative totals of relevant articles produced, he produced a straight line.[4] On the basis of these results Bradford formulated a mathematical model to describe "reference scattering". A number of authors have used his method to analyze the characteristics of literature in any particular subject field.

Cole studied the needs of a narrow field of a petroleum industry.[5] He designed a yardstick for assessing the efficiency correlation between the holdings of this library and its users. In a closely related article, Cole established that in the field of petroleum literature there is a definite relationship between usage and age of journal.[6] The relationship is expressed by him as follows:

a) Log per cent usage of journals older than X years = $2 - K_x$, where K is constant for any particular collection.

b) Demand upon the collection is characterized by its median age, $x(\frac{1}{2}) = 0.301/K$

These relationships have obvious potential applications of establishing an efficient retention pattern in a given library. In applying these principles to the industrial library with which Cole was associated, it was found that 84 per cent of all articles read were published within the last six years.

[4]Bradford, S.C. - "Sources of Information of Specific Subjects", Engineering 137, No. 3550:85-6, January 1934

[5]Cole, P.F. - "A Look at Reference Scattering", Journal of Documentation, 18:58-64, June 1962

[6]Cole, P.F. - "Journal Usage Versus Age of Journal", Journal of Documentation, 19:1-11, March 1963

Cole's studies, limited to the petroleum industry, seem to bear some similarities to a study much broader in scope. Garfield analyzed the citations that appear in some 2200 scientific journals.[7] He found that a small core of about 152 journals accounted for approximately half of all citations, and about 500 scientific journals published around 70 per cent of all articles. Garfield has confirmed Bradford's findings that journals can be ranked by frequency of being cited.

Strain reports the result of research done to develop a formula that would help to identify the most suitable journals for storage.[8] Strain analyzed serial circulation records and established the percentage of collection used, as well as the most frequently requested titles and their age distribution.

Trueswell proposes reducing the primary access collection based on the experience of user needs[9]. He, too, employs the circulation record technique, wherein the last circulation date of the book is examined as a parameter of user requirements. He suggests that research libraries using this design could reduce their collections by 60 to 70 per cent and still fill 99 per cent of user needs. His assumptions have yet to be tested in a broad-scale basis.

[7]Garfield, Eugene - "Citation Analysis as a Tool in Journal Evaluation", Science, 178:471-479, November 1972

[8]Strain, Paula M. - "A Study of the Usage and Retention of Technical Periodicals", Library Resources & Technical Services, 10:295-304, Summer 1966

[9]Trueswell, Richard W. - "A Quantitative Measure of User Circulation Requirements and its Possible Effect on Stack Thinning and Multiple Copy Determination", American Documentation 16, 20-25, January 1965

Probably one of the most thorough experimental designs depicting the patrons' use of books in a large university library was executed by Fussler and Simon.[10] The authors took a random sample of the Economics and Teutonic Languages collections of the University of Chicago, which was then weighted to eliminate the heavy representation of modern works. This sample of books was examined for recorded past use during 1954 - 1959. Fussler and Simon suggested that the best predictor for future use was past use of the title; in fact, the longer the observation period, the better. They theorized that books owned by a library but seldom used have far less probable future demand than books with continuous use records.

The authors were concerned with the accuracy of prediction of two quantities:

1) The portion of titles that will actually go to storage under a rule that attempts to send a given portion.

2) The number of titles that will be recalled from storage from the same group.

They have suggested the following formula to be applied to both predictions at the 95 per cent confidence level:

$$SD = \sqrt{\frac{pg}{n}}$$

where p = the sample portion sent to storage

g = sample portion not sent to storage

n = sample size

SD = standard deviation

[10] Fussler, Herman and Simon, Julian - "Patterns in the Use of Books in Large Research Libraries", Chicago, University of Chicago Press, 1969

The research design developed by Fussler and Simon has great potential in the research design to be developed for other academic libraries. It should be noted, however, that their design was limited to Economics and Teutonic Languages. Other libraries will have to cover all subject areas of curricular interest on their campuses. Further, the University of Chicago study, which concerned itself with a unique research library situation, cannot be applied to a smaller institutional group without some modification and possible simplification.

The lack of ready-made data available and applicable to small and medium-sized libraries reaffirms the need to conduct a pilot study before universities make any further commitments to future library developments.

IV. METHODOLOGICAL and SAMPLING TECHNIQUES of THE STUDY

The techniques employed in this study are well established for most disciplines; however, their application to library situations is somewhat new. The studies examined during the literature search revealed the need for further experiments conducted under different conditions and institutional settings.

The techniques proposed are being tested at Florida Atlantic University in a scaled-down model program. If the FAU pilot project proves to be valid, it could be a desirable method for other library situations.

As stated earlier, the central purpose of this study is to design and test a method by which a medium-sized academic library can determine the relationship between usage and age of material. The Circulation Record will be the basis for this examination.

A general statement of objective of this study as it relates to the storage decision problem is expressed as follows:

> Given a collection of books of known size and a decision to select some of the books for storage in a remote area, determine a selection rule which will select books in such a manner that less than five per cent of those books selected for storage will be circulated during any given year.

The following hypotheses support the testing of one or more variables to determine if they influence the frequency-of-use pattern. These hypotheses will be compared for results. It is hoped that such comparisons will justify the complex statistical procedures required by the use of multi-variables.

The hypotheses to be tested are:

1) H_0 There is no significant difference of aging process* (obsolescence as demonstrated by decreasing use) among the 23 major subject areas as defined by L.C. Classification schedule.**

2) H_0 There is no significant difference of aging process* (obsolescence as demonstrated by decreasing use) among the 23 major subject areas as defined by L.C. Classification schedule** when U.S. publications are measured against publications of foreign origin.

3) H_0 There is no significant difference of aging process* (obsolescence as demonstrated by decreasing use) among the 23 major subj. areas as defined by L.C. Classification schedule** when English language publications are measured against non-English language publications.

*Age determined by publication date
** For L.C. Classes, see Appendix A

Assumptions:

1) At any given moment each book in the library has a random probability of being used within some specified period of time.

2) The larger the number of volumes in a given subject field, the smaller the probability that a given volume will be used.

3) It is impossible to arrive at any meaningful statement looking at the use history of a single book; therefore, it is necessary to group books together on the basis of common characteristics.

4) Recorded circulation of books is a reasonable index of all use of material.

5) Books that develop little recorded use develop little browsing.

6) Within broad subject areas and types of material use is a suitable criteria for sorting materials into different levels of accessibility.

7) As books grow older they are of progressively less interest to readers.

8) Frequency of use is a reasonable index of the value of a book to an academic community.

9) Economic factors may make it highly desirable to segregate books on the basis of their value and into two or more levels of accessibility.

Exclusions:

Several types of library materials have been excluded from the study. These include maps, periodicals, serials, newspapers, microtexts, and material that has been selected for the Reference Collection. These materials represent a rather unique problem of

their own and, consequently, they should not be lumped into the general circulating book collection category when making a study of use problems.

Also excluded are books published after January 1, 1966. These are assumed to be in a heavy demand category.

The Sampling Unit:

The sampling unit will consist of the monographic collection of the library. The "title" will be the basic unit for the sampling, because it is easier to define than "copy", "volume" or "edition". The title will be the basic unit for the analysis, also.

The Sampling Technique:

The basic approach to this study is the sampling of titles of books listed in the shelflist.

A shelflist is an inventory of library holdings in which the cards are arranged by their classification number. It reflects the actual shelf arrangement of books. Drawing a sample from the shelflist is conceptually the same as taking a sample of titles from the shelves if all the library's holdings were on the shelves.

The definition of universes includes only the monographic titles in each subject area.

The object of sampling is to allow each qualified title within the 23 subject areas the same probability of entering the sample.

The conceptual approach to fix the sample size is necessary to minimize the cost of sampling. Two hundred titles in each subject area will serve as an adequate sample size for each L.C. number category.

The stratified random sampling technique will be employed to accomplish this. The stratified sample plan is as follows:

Publication Date	
Period	Sample
Pre-1900	take all
1901 - 1925	take 4 of every 5 titles
1926 - 1935	take 3 of every 5 titles
1936 - 1945	take 2 of every 5 titles
1946 - 1955	take 1 of every 5 titles
1956 - 1965	take 1 of every 10 titles

It is conceivable that in some subject areas the library will have large holdings resulting in more than 200 titles after the first stratified random sampling is finished. In such cases, the sample size will be further reduced at random within each group to maintain the balance among age groups. This should result in a true stratified sample of 200 cards.

Handling the Sample, Including the Experimental and Control Groups

When the composite sample of the 23 subject groups has been established, a proportional random sample of the prime sample will be removed from the shelves and locked in a room to simulate a conceptualized storage facility. Special attention will be given to dividing the 200 titles so that each age group will be proportionately represented.

The portion remaining on the shelves will be called the Control Group.

The portion locked up will be our Experimental Group.

The 4600 titles selected will be keypunched to enable monitoring transactions through our computerized circulation system for three months, repeated four times. Thus, circulation activities for both Experimental and Control groups will be cumulatively recorded for the benefit of this project for a total of twelve months.

Bibliographic and Use Control

Use of the entire sample will be controlled through the University Computer Center. Daily circulation records will be checked by L.C. number and any use will be noted.

The Public Card Catalog will be marked for the Experimental Group locked in storage. The card of each Experimental title will be identified as being available within 24 hours through application to the staff. The Control Group will not bear any mark which would indicate they are under experimental study.

V. STATISTICAL METHODOLOGY

The total observation period will extend through one year. Three-month observation cycles will be duplicated four times to obtain more than one date per cell.

Our investigation will consider "publication date", "language" and "country of origin" as independent variables. The measured circulation uses of the test samples will be the dependent variables.

The following variables will be examined to establish which combination of functions will succeed best under operations in predicting the future level of use of monographic titles in the collection, where:

Publication Date	=	A
Language	=	B
Country of Origin	=	C

We will investigate the following main effects and interaction effects through the analysis of variances:

		A_1	A_2	A_3	A_4	A_5	A_6
C_1	B_1	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
	B_2	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
C_2	B_1	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
	B_2	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4

In this paradigm the publication date coincides with the six age periods used for the stratified sample plan:

A_1 = pre- 1900

A_2 = 1901 - 1925

A_3 = 1926 - 1935

A_4 = 1936 - 1945

A_5 = 1946 - 1955

A_6 = 1956 - 1965

The distinction between English and all other languages will be made thus:

B_1 = English language

B_2 = non-English languages

Country of origin will attempt to establish the level of influence this has on the reader. Again, we will use a two-tiered method, i.e.:

C_1 = U.S. publications

C_2 = All non-U.S. publications

Based on the above, we can formulate the following:

Analysis of Variance

A	(a-1)
B	(b-1)
C	(c-1)
AB	(a-1) (b-1)
AC	(a-1) (c-1)
BC	(b-1) (c-1)
ABC	(a-1) (b-1) (c-1)
S/ABC	abc(n-1)

We will assume there will be a significant F-ratio. The particular statistical test yielding the answer is:

$$F = \frac{\text{Between group variance}}{\text{Within group variance}}$$

It is expected that the use of variables of publication date, language and country of origin will be the most powerful design and will produce the most reliable answers.

These analyses will be used for each of the 23 L.C. subject classes examined.

VI. SUMMARY AND RECOMMENDATIONS

The growth of large and medium-sized university libraries and the high cost of library buildings have led university administrators to consider the possibility of less expensive housing for lesser used books. Shelving such books in a warehouse could reduce housing costs considerably.

The possibility of storage of books raises these questions:

What percentage of its books should the library place in a second-level-of-access facility?

What will be the public reaction to such a measure?

Will the public accept the concept of - for example - a 24-hour delivery system?

The answers to these questions are speculative at best. They will depend greatly upon the choices made by administrators who, no doubt, will select solutions that promise to maximize the welfare of the members of the university.

This presentation depended on the following concepts:

The recorded circulation use of books is a reasonable index for all use for material.

As books grow older they become progressively less valuable to the reader.

If recorded use is a satisfactory index of value, one can divide the collection into high and low value sections.

The study assumes there will be a significant difference in the aging process of books among disciplines. This difference will probably increase when examined in correlation with language and country of origin.

This study assumes many generalizations that imply even further generalizations that lead to necessary over-simplification. For example, it was stated that the function of age will be accepted as an effective variable; yet, we could argue that this phenomenon is probably greater in disciplines where knowledge is hierarchical as opposed to historical or cumulative disciplines. It is hoped that the pilot study will bear out the validity of our assumptions.

Further study is needed to expand this investigation to cover the questions raised above. It is specifically suggested that a regression study be conducted, using the best results that emerged from the Analysis of Variance data discussed earlier. The regression study may establish the following with a desired level of accuracy of prediction:

1) The portion of titles that will actually be sent to storage under a given rule that attempts to send a given portion.

2) The number of titles that will be withdrawn from a given group sent to storage.

The given proportion will be arbitrary and will be set administratively. It will be a percentage of those books below the cutting point of less than five per cent use per age group, also correlated with language and country of origin.

To illustrate the procedure for a publication date plus language rule, the 95 per cent confidence limit, and 25 per cent to storage estimate, the formula would be:

$$2\frac{.25 \times .75}{N} = \text{the number sent to storage}$$

where N = the sample size.

The second prediction - i.e., the accuracy of prediction of number of withdrawals in a given number of books sent to storage - may be approached in terms of binomial confidence limits.

The validity of this prediction method has been tested by Fussler and Simon at the University of Chicago as discussed earlier. However, its application in a medium-sized institution needs to be confirmed.

* * * * * * * * * *

APPENDIX A

List of Library of Congress Classification

A	General Works
B	Philosophy
C	History - Auxiliary Sciences
D	History and Topography
E	History - U. S.
F	History - America except U.S.
G	Geography - Anthropology
H	Social Sciences
J	Political Science
K	Law
L	Education
M	Music
N	Fine Arts
P	Language and Literature
Q	Science
R	Medicine
S	Agriculture
T	Technology
U	Military Science
V	Naval Science
Z	Bibliography, Library Science